Journal of
Cheminformatics

ORAL PRESENTATION   **Open Access**

A Branch-and-Bound approach for tautomer enumeration

Torsten Thalheim[*], R-U Ebert, R Kühne, G Schüürmann

For molecules with mobile H atoms, the result of quantitative structure-activity relationships (QSARs) depends on the position of the respective hydrogen atoms. Thus to obtain reliable results, tautomerism needs to be taken into account.

In the last years, many approaches were introduced to achieve this. In this work we present a further development of our previous algorithm based on InChI-layers. While the InChI ansatz supports only heteroatom-tautomerism, we suggest an extension regarding carbon atoms too. Whereas with other tautomer generating algorithms the hydrogen shifts are based on pattern-rules, we try to overcome the rule constriction and evolve a more common solution. The advantage of our approach is quite simple. Due to the avoidance of a rule system with its necessity for exceptions to the rules, we can apply our solution to any kind of tautomerism definition.

We set up a Branch-and-Bound approach, which is optimized to generate a complete enumeration of all tautomers, with regard to a certain definition, from any structure. With few and easy decisions like symmetry detection, we avoid a lot of calculation overhead. Decisions with significant influence on the algorithm efficiency are made as early as possible.

We have set up several kinds of tautomer definitions and derived a stable definition covering the major kinds of protropic tautomerism. Furthermore we analyzed, what expenditure of time for large databases (case study: more than 70,000 entries) is needed to investigate which structures have tautomers and which not for more than 99% of the database entries.

UFZ Department Ecological Chemistry - Helmholtz Centre for Environmental Research, Permoser Str. 15, 04318 Leipzig, Germany